

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
26 July 2001 (26.07.2001)

PCT

(10) International Publication Number
WO 01/53455 A2

(51) International Patent Classification⁷: C12N

[US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US).
LIU, Chenghua [CN/US]; 1125 Ranchero Way #14, San
Jose, CA 95117 (US). **DRMANAC, Radoje, T.** [YU/US];
850 East Greenwich Place, Palo Alto, CA 94303 (US).

(21) International Application Number: PCT/US00/35017

(22) International Filing Date:

22 December 2000 (22.12.2000)

(74) Agent: **ELRIFI, Ivor, R.**; Mintz, Levin, Cohn, Ferris,
Glovsky, and Popeo, P.C., One Financial Center, Boston,
MA 02111 (US).

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

09/471,275	23 December 1999 (23.12.1999)	US
09/488,725	21 January 2000 (21.01.2000)	US
09/552,317	25 April 2000 (25.04.2000)	US

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,
DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,
TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(63) Related by continuation (CON) or continuation-in-part
(CIP) to earlier applications:

US	09/488,725 (CIP)
Filed on	21 January 2000 (21.01.2000)
US	09/596,196 (CIP)
Filed on	17 June 2000 (17.06.2000)
US	09/653,274 (CIP)
Filed on	31 August 2000 (31.08.2000)

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished
upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(71) Applicant (*for all designated States except US*): **HYSEQ,
INC.** [US/US]; 670 Almanor Avenue, Sunnyvale, CA
94086 (US).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **TANG, Y., Tom**

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and
uses thereof.

WO 01/53455 A2

WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-739, a mature protein coding portion of SEQ ID NO: 1-739, an active domain of SEQ ID NO: 1-739, and complementary sequences thereof.
2. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide hybridizes to the polynucleotide of claim 1 under stringent hybridization conditions.
3. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide has greater than about 90% sequence identity with the polynucleotide of claim 1.
4. The polynucleotide of claim 1 wherein said polynucleotide is DNA.
5. An isolated polynucleotide of claim 1 wherein said polynucleotide comprises the complementary sequences.
6. A vector comprising the polynucleotide of claim 1.
7. An expression vector comprising the polynucleotide of claim 1.
8. A host cell genetically engineered to comprise the polynucleotide of claim 1.
9. A host cell genetically engineered to comprise the polynucleotide of claim 1 operatively associated with a regulatory sequence that modulates expression of the polynucleotide in the host cell.
10. An isolated polypeptide, wherein the polypeptide is selected from the group consisting of:

- (a) a polypeptide encoded by any one of the polynucleotides of claim 1; and
 - (b) a polypeptide encoded by a polynucleotide hybridizing under stringent conditions with any one of SEQ ID NO:1-739.
11. A composition comprising the polypeptide of claim 10 and a carrier.
12. An antibody directed against the polypeptide of claim 10.
13. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
- a) contacting the sample with a compound that binds to and forms a complex with the polynucleotide of claim 1 for a period sufficient to form the complex; and
 - b) detecting the complex, so that if a complex is detected, the polynucleotide of claim 1 is detected.
14. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
- a) contacting the sample under stringent hybridization conditions with nucleic acid primers that anneal to the polynucleotide of claim 1 under such conditions;
 - b) amplifying a product comprising at least a portion of the polynucleotide of claim 1; and
 - c) detecting said product and thereby the polynucleotide of claim 1 in the sample.
15. The method of claim 14, wherein the polynucleotide is an RNA molecule and the method further comprises reverse transcribing an annealed RNA molecule into a cDNA polynucleotide.
16. A method for detecting the polypeptide of claim 10 in a sample, comprising:

- a) contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex; and
- b) detecting formation of the complex, so that if a complex formation is detected, the polypeptide of claim 10 is detected.

17. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

- a) contacting the compound with the polypeptide of claim 10 under conditions sufficient to form a polypeptide/compound complex; and
- b) detecting the complex, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.

18. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

- a) contacting the compound with the polypeptide of claim 10, in a cell, under conditions sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and
- b) detecting the complex by detecting reporter gene sequence expression, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.

19. A method of producing the polypeptide of claim 10, comprising,

- a) culturing a host cell comprising a polynucleotide sequence selected from the group consisting of a polynucleotide sequence of SEQ ID NO: 1-739, a mature protein coding portion of SEQ ID NO: 1-739, an active domain of SEQ ID NO: 1-739, complementary sequences thereof and a polynucleotide sequence hybridizing under stringent conditions to SEQ ID NO: 1-739, under conditions sufficient to express the polypeptide in said cell; and
- b) isolating the polypeptide from the cell culture or cells of step (a).

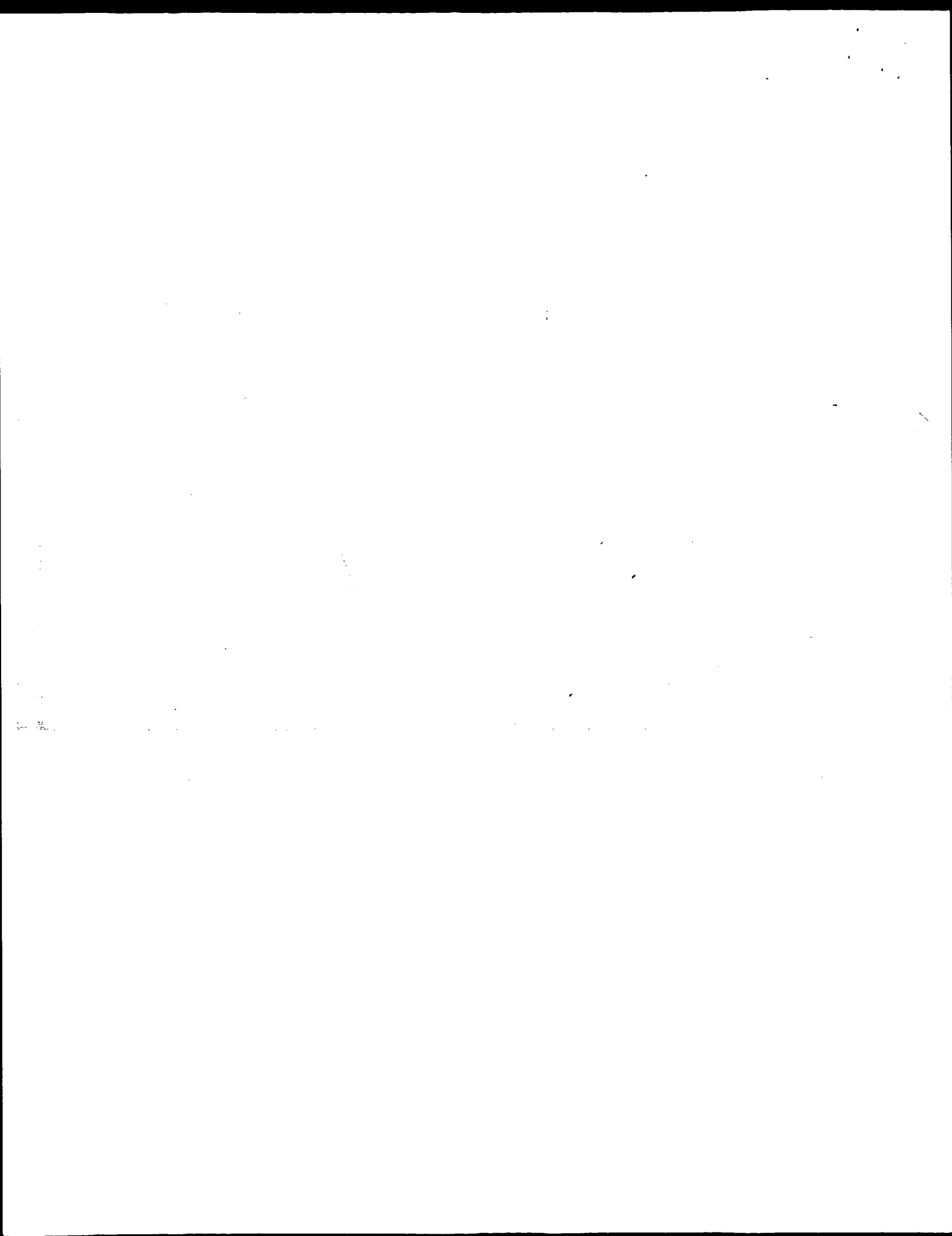
20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 740-1478, the mature protein portion thereof, or the active domain thereof.
21. The polypeptide of claim 20 wherein the polypeptide is provided on a polypeptide array.
22. A collection of polynucleotides, wherein the collection comprises the sequence information of at least one of SEQ ID NO: 1-739.
23. The collection of claim 22, wherein the collection is provided on a nucleic acid array.
24. The collection of claim 23, wherein the array detects full-matches to any one of the polynucleotides in the collection.
25. The collection of claim 23, wherein the array detects mismatches to any one of the polynucleotides in the collection.
26. The collection of claim 22, wherein the collection is provided in a computer-readable format.
27. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.
28. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising an antibody that specifically binds to a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

<210> 713
 <211> 524
 <212> DNA
 <213> Homo sapiens

<400> 713
 atccccacag ggtaatgggt gtcccgatgt caccgggggac tctgtgatcc gtgttccoct 60
 gaccctccta gtgcacaact tggccggggt cactggggctc ctgcaccact gectgtcagg 120
 tccgtgcca gccccaagcc cccaccagc catgagctcc tccagaaagg accacctcgg 180
 cgccagcagc tcagagcccc tcccggtcat cattgtgggt aacggccct ctggtatctg 240
 cctgtcctac ctgctctccg gctacacacc ctacacgaag ccagatgcca tccaccaca 300
 cccctgctg cagaggaagc tcaccgaggc cccgggggtc tccatcctgg accaggacct 360
 ggactacctg tccgaaggcc tcgaaggccg atcccaaagc cccgtggccc tgctctttga 420
 tgcccttcta cgccagaca cagactttgg gggaaacatg aagtcggtcc tcacctggaa 480
 gcaccggaag gagcacgcca tccccacgt ggttctgggc cgga 524

<210> 714
 <211> 2468
 <212> DNA
 <213> Homo sapiens

<400> 714
 gaatcgacgc acgctgctgc agcgtgcca gctggaagg agctgccccg cgccggaggga 60
 ggaagttagg cccgggaccg ccaggccacc accggccgccc tcagccatgg acgctccct 120
 ggagaagata gcagacccca cgttagctga aatgggaaaa aacttgaagg aggcagtga 180
 gatgctggag gacagtccga gaagaacaga agaggaaaat ggaaagaagc tcatatccgg 240
 agatattcca ggcccactcc agggcagtg gcaagatag gtgagcatcc tccagttagt 300
 tcagaatctc atgcatggag atgaagatga ggagccccag agccccagaa tccaaaatat 360
 tggagaacaa ggtcatatgg ctttgttggg acatagtctg ggagcttata tttcaactct 420
 ggacaaagag aagctgagaa aacttacaac taggatactt tcagatacca ccttatggct 480
 atgcagaatt ttcagatatg aaaatgggtg tgcttatttc cacgaagagg aaagagaagg 540
 acttgcaaaag atatgtaggc ttgccattca ttctcgatat gaagacttcg tagtggatgg 600
 cttcaatgtg ttatataaca agaagcctgt catatatctt agtctgtctg ctagacctgg 660
 cctgggccaac tacctttgtat atcagctcgg cttgcccttc cctgtcttgt gccgtgtacc 720
 ctgtaacact gtgtttggat cccagcatca gatggatgtt gccttccttg agaaactgat 780
 taaagatgat atagagcgag gaagactgcc cctgttgctt gtcgcaaatg caggaaaggc 840
 agcagttaga cacacagaca agattgggag attgaaagaa ctctgtgagc agtatggcat 900
 atggcttcat gtggaggggt tgaatctggc aacattggct ctgggttatg tctcctcatc 960
 agtgcctggc gcagccaaat gtgatagcat gacgatgact cctggccctg ggcctgggtt 1020
 gccagctgtt cctgcggtga cactgtataa acacgatgac cctgccttga ctttagttgc 1080
 tggctcttaca tcaaataaag ccacagacaa actccgtgcc ctgcctctgt gggtatcttt 1140
 acaatacttg ggacttgatg ggtttgtgga gaggatcaag catgcctgtc aactgagtca 1200
 acggttgacg gaaagtgtga agaaagtga ttacatcaaa atcttggttg aagatgagct 1260
 cagctcccca gtggtgggtg tcagattttt ccaggaatta ccaggctcag atccggtgtt 1320
 taaagccgtc ccagtgcaca acatgacacc ttccaggagtc ggccgggaga ggccagctgt 1380
 tgacgcgctg aatcgctggc tgggagaaca gctgaagcag ctgggtgcctg caagcggcct 1440
 cacagtcatg gatctggaag ctgagggcac gtgtttgctg ttcagccctt tgatgacgcg 1500
 agcaggtaaa ccaggcttgg tggacatccc ttgcttttgt tctggggctg ctgggttagat 1560
 tagcttgccc ttatgatact ccattctcct agagttatta gcagctcttt ttggaggggc 1620



atcttctctt	cttttgggct	aaatttaggt	agattagcat	tcccatgtaa	cttaccagaa	1680
tcagaatgag	aattcagaag	tcacctgaat	tggccgggca	tggtgggtca	cacctgtaat	1740
cccagcacct	tgaggaggcca	aggcaggcag	atcatctgag	gtcaggagtt	cgagaccagc	1800
ctggccaaca	tagtgaaatc	cgcgccctac	taaaaatata	aaaaatttagc	caggcaccct	1860
gtccacagcc	cccacacaga	ctcgaggggc	ccccatctcc	tggtctgaac	ccaacagggt	1920
ggtcccactg	tgggaccaca	accaggatatg	actgtgtgag	aagcaggctc	actaccaggc	1980
taccagggag	cacaggggag	caggcgccac	cttgaggcat	aaaccagag	aaacaagacc	2040
tccaagacgg	ccaggcactg	gggcacacgc	cggtaacaca	gcaccgtggg	agctgagacg	2100
gaaggatcgc	ctgagcccag	gattttgaaa	ccaccctggg	caacacagtg	agaccccgtg	2160
tctacaaaaa	aatacacatt	agccaggcat	ggcggcatgc	gcctggggtc	ccaagtactc	2220
gggaggtaga	ggagagaaaa	atcacttgag	cccagagagg	tcaaggctac	agggagctga	2280
gatcgcatca	ctgtactcca	gctgggggtga	aacggcgaga	ctctacctca	aaaataaata	2340
aatacatata	taattaataa	ataaaacatc	aaagaccagc	cgacctaaact	ccatctaaaa	2400
tacacaaactt	ctacgcaaaa	tataaataaa	attagaaaac	aaactacaat	ctcagaaaag	2460
cactagca						2468

<210> 715
 <211> 924
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(924)
 <223> n = a, t, c or g

<400> 715						
tttcgtgtaa	gatataactc	aactttgaaa	atgtcagccg	ttatagttga	agaaatctga	60
cccaagagac	ttcgctccgc	tgcaagatgg	aagggaagctt	aagtaagaca	taaatttgta	120
atgaacttgc	tcacaacatc	cgccgccact	gtgacttgca	gtcatcatcc	attaccacaa	180
aattagttgc	aggatggcta	ctcgatccc	tccacacatg	atcatcagta	tttgctcct	240
gtgtcccaac	eggcctgagt	caagggttacg	actcaactgat	taaaaagagg	gactttttca	300
aatactttgc	actttttgatt	gtgtattatg	gataccaagg	aagagaagaa	ggaacggaaa	360
caaagttatt	ttgctcgact	gaaaaagaaa	aaacaagcca	aaacaaatgc	agagacagcc	420
tcagctgtag	ctacaaggac	tcatactggg	aagggaagata	ataatacagt	agtttttagag	480
ccagacaagt	gcaacattgc	tgtggaagag	gaatatatga	ctgatgagaa	aaaaaagaga	540
aaaagtaatc	agttaaagga	gatcaggcgt	acagaactaa	agagatatta	tagtattgat	600
gacaatcaaa	acaaaacaca	tgataaaaaa	gagaagaaga	tggtgggtca	gaagcccat	660
gggactatgg	aatacactgc	tggaaaccag	gacacccata	actccatagc	actgaaattt	720
aacatcactc	ccaataaatt	ggtggaactg	aataaaacttt	tcacacatac	tattgttcca	780
ggccagggtcc	tttttgtgcc	agatgccaac	tctccttcca	gtaccttaag	gctatcatca	840
tccagtcctg	gtgctactgt	ctctccttca	tcatnagatg	cagaatatgn	taattggctg	900
atgctgactt	agcacgggaag	gctt				924

<210> 716
 <211> 679
 <212> DNA
 <213> Homo sapiens

35 40 45
 Pro Ala Met Ser Ser Ser Arg Lys Asp His Leu Gly Ala Ser Ser Ser
 50 55 60
 Glu Pro Leu Pro Val Ile Val Gly Asn Gly Pro Ser Gly Ile Cys
 65 70 75 80
 Leu Ser Tyr Leu Leu Ser Gly Tyr Thr Pro Tyr Thr Lys Pro Asp Ala
 85 90 95
 Ile His Pro His Pro Leu Leu Gln Arg Lys Leu Thr Glu Ala Pro Gly
 100 105 110
 Val Ser Ile Leu Asp Gln Asp Leu Asp Tyr Leu Ser Glu Gly Leu Glu
 115 120 125
 Gly Arg Ser Gln Ser Pro Val Ala Leu Leu Phe Asp Ala Leu Leu Arg
 130 135 140
 Pro Asp Thr Asp Phe Gly Gly Asn Met Lys Ser Val Leu Thr Trp Lys
 145 150 155 160
 His Arg Lys Glu His Ala Ile Pro His Val Val Leu Gly Arg
 165 170 174

<210> 1453
 <211> 518
 <212> Amino acid
 <213> Homo sapiens

<400> 1453
 Asn Arg Arg Thr Arg Ala Gln Arg Cys Gln Arg Gly Arg Ser Cys Gly
 1 5 10 15
 Ala Arg Glu Glu Glu Val Glu Pro Gly Thr Ala Arg Pro Pro Ala
 20 25 30
 Ala Ser Ala Met Asp Ala Ser Leu Glu Lys Ile Ala Asp Pro Thr Leu
 35 40 45
 Ala Glu Met Gly Lys Asn Leu Lys Glu Ala Val Lys Met Leu Glu Asp
 50 55 60
 Ser Gln Arg Arg Thr Glu Glu Glu Asn Gly Lys Lys Leu Ile Ser Gly
 65 70 75 80
 Asp Ile Pro Gly Pro Leu Gln Gly Ser Gly Gln Asp Met Val Ser Ile
 85 90 95
 Leu Gln Leu Val Gln Asn Leu Met His Gly Asp Glu Asp Glu Glu Pro
 100 105 110
 Gln Ser Pro Arg Ile Gln Asn Ile Gly Glu Gln Gly His Met Ala Leu
 115 120 125
 Leu Gly His Ser Leu Gly Ala Tyr Ile Ser Thr Leu Asp Lys Glu Lys
 130 135 140
 Leu Arg Lys Leu Thr Thr Arg Ile Leu Ser Asp Thr Thr Leu Trp Leu
 145 150 155 160
 Cys Arg Ile Phe Arg Tyr Glu Asn Gly Cys Ala Tyr Phe His Glu Glu
 165 170 175
 Glu Arg Glu Gly Leu Ala Lys Ile Cys Arg Leu Ala Ile His Ser Arg
 180 185 190
 Tyr Glu Asp Phe Val Val Asp Gly Phe Asn Val Leu Tyr Asn Lys Lys
 195 200 205
 Pro Val Ile Tyr Leu Ser Ala Ala Ala Arg Pro Gly Leu Gly Gln Tyr
 210 215 220
 Leu Cys Asn Gln Leu Gly Leu Pro Phe Pro Cys Leu Cys Arg Val Pro
 225 230 235 240
 Cys Asn Thr Val Phe Gly Ser Gln His Gln Met Asp Val Ala Phe Leu
 245 250 255
 Glu Lys Leu Ile Lys Asp Asp Ile Glu Arg Gly Arg Leu Pro Leu Leu
 260 265 270
 Leu Val Ala Asn Ala Gly Thr Ala Ala Val Gly His Thr Asp Lys Ile

